Homework Questions

Due date month 01, 2018 13:30 EET

Q:(10pts) Create your own expressions that calculate *variance* and *standart deviation* of an arbitrary `x` vector. Compare your results by *R*'s `var` and `sd` functions.

Q:(10pts) `x` is an integer vector from `2` to `5000` by interval of `3`. Calculate *variance* and *standart deviation* of remainders of `x` division by `13`.

Q:(10pts) $f(x) = 2x^6 - 13x^5 + 26x^4 - 7x^3 - 28x^2 + 20x$ and `x` is a numeric vector from `-1` to `3` by interval of `0.05`. Calculate *variance* and *standart deviation* of `f(x)` and plot `f(x)`.

Q:(10pts) A dependent function chain is defined as $h(x) = \frac{log(x)-1}{\sqrt{x}}$, $g(x) = e^{\sqrt{h(x)}}$ and $f(x) = \sin g(x)^{\cos g(x)}$. If `x` is an integer array in the interval of `[4, 250]`,

- Calculate *standart deviation* and *variance* of $f(x)$.
- Calculate *minimum*, *1st quartile*, *median*, *mean*, *3th quartile* and *maximum* of f(x).

Q:(20pts) You have a two air quality stations which first is located in the city center of Bursa and the latter one is at Uludağ. Their heights from mean sea level are $155m$ and $2543m$; and PM_{10} concentrations are $87\mu g/m^3$ and $23\mu g/m^3$, respectively. Assuming PM_{10} concentration change by height is linear, what is the PM_{10} concentration at a village that height is $240m$?

Q:(20pts) You are an English cryptology expert in World War II and MI6 agents captured a cryptred German telegraph communication. You figured out that each letter is shifted by `2` to right in crypted text. So, write the decryption algorithm in R language and help to end the war. You can test your algorithm by the piece of text below. (*Use only English alphabet to decrypt*)

```
crypted_text <- c("y", "k", "n", "n", "m", "q", "o", "o", "g", "p")
```

Q:(20pts) In numerical integration,

$$\int_a^b f(x)dx \approx \frac{b-a}{n}\sum_{i=1}^{n+1} f(x_i)$$

is known as mid-point rule. According to the definition, calculate the $\int_0^2 sin(x)dx$ integral.

Hint: Divide the interval [a, b] into `n` subintervals of equal width. Theoretical solution is,

$$\int_a^b sin(x)dx = cos(a) - cos(b)$$

Q4:(20pts) Taylor series expansion of `sin(x)` is,

$$sin(x) = \sum_{n=0}^{k=\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \approx x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \dots$$

Using the formula above, write the code that calculates the value of $sin(0.5)$ for `k = 5`.

Q:(10pts) Write your own R question and the answer. What kind of a question would be useful for someone that just begins?

Q:(10pts) Create the table below as a `data.frame` object;

name	order	is.weekend
Monday	1	FALSE
Tuesday	2	FALSE
Wednesday	3	FALSE
Thursday	4	FALSE
Friday	5	FALSE
Saturday	6	TRUE
Sunday	7	TRUE

Q:(20pts) Create a numeric matrix with 5 columns and 15 rows using following `x` vector.

```
set.seed(1); x <- rnorm(75)
```

1. What is the *mean* of each columns?
2. What is the *mean* of each rows?
3. What are the *row* and *column* indices of **minimum value** in the *matrix*?
4. What are the *row* and *column* indices of **maximum value** in the *matrix*?

Q:(20pts) Create a `list` having the properties below and discuss the struct of the resultant object.

1. First element is vector from 1 to 10.
2. Second element is *names of persons* in the class.
3. Third element is the `matrix` created in first question.
4. Fourth element is the `data.frame` created in second question.

Q:(10pts) `airquality` data frame is one of datasets comes with base R (help: `?airquality`) and is consist of daily air quality measurements of *ozone*, *solar radiation*, *wind speed* and *temperature*.

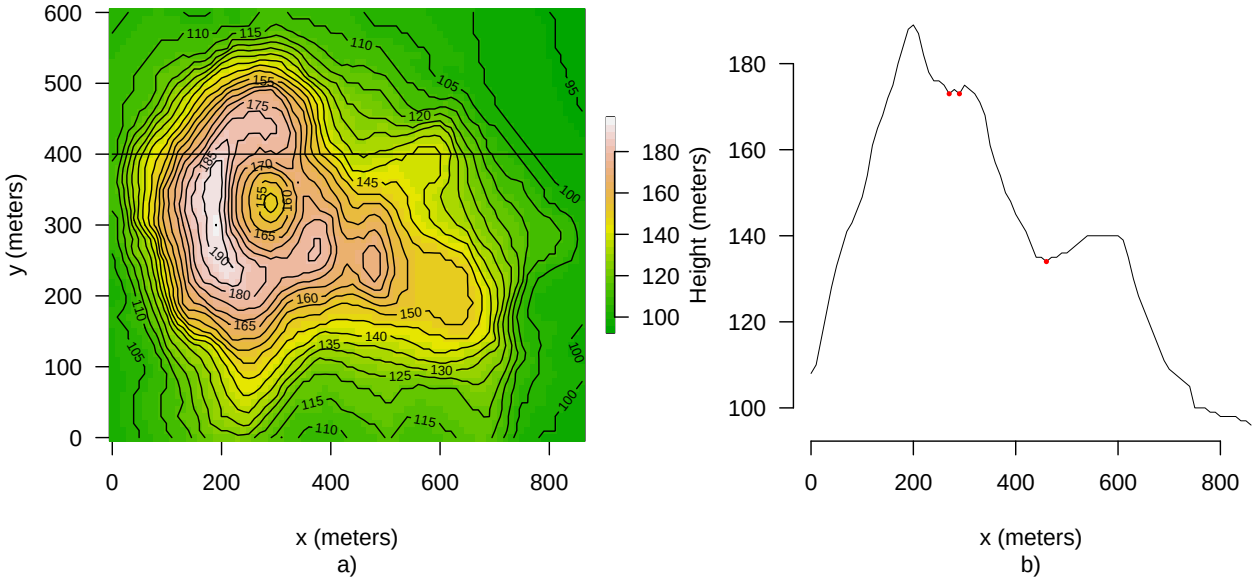
1. What are the *minimum*, *1st quartile*, *median*, *mean*, *3th quartile* and *maximum* of the measurements.
2. Calculate *standart deviation* and *variance* of the measurements.

Q:(40pts) As you see in previous question, *ozone* data has `NA` values. Calculate the *mean* of *ozone* by applying each of the methods below and conclude which one is better? Or each of them has their own advanteges?

1. Mean of *ozone* by removing `NA`s. *(Already calculated in previous question)*
2. **(15pts)** Fill NAs in *ozone* by replacing NA by mean of *ozone* at step 1 then calculate *mean*.
3. **(15pts)** Fill NAs in *ozone* by linear interpolation then calculate *mean*. (**Hint:** `approx` function)

Maunga Whau (Mt Eden) is one of about 50 volcanos in the Auckland volcanic field. This data set (`?volcano`) gives topographic information for Maunga Whau on a 10m by 10m grid. *Figure a)* shows contour plot of the volcano Maunga Whau. *Figure b)* shows the cross-section of the volcano at `y = 400m`. Answer the following questions in the light of this information.

```
z <- volcano # Store volcano data in z
x <- (0:(nrow(z) - 1)) * 10 # x-values
y <- (0:(ncol(z) - 1)) * 10 # y-values
y400 <- z[,41] # heights at y = 400m
```



Q1:(10pts) Answer the questions below using R;

1. What is the *maximum height* of the volcano?
2. What is the `[x, y]` coordinate of maximum height?
3. What is the *minimum height* of the volcano?
4. What is the `[x, y]` coordinate of minimum height?

Q2:(30pts) Data at 6 cells in the matrix (`z`) was removed becuse of quality assurence *(run the command below)*.

```
z <- volcano # Store volcano data in z
set.seed(1)
indices <- round(runif(6, 1, length(z))) # indices to set NA
z[indices] <- NA # set to NA in purpose
```

1. Find the `x` and `y` coordinates of `NA`s.
2. Fill the `NA`s by mean of the 8 values surrounding `NA` cell.

Q3:(30pts) *Figure b)* shows the cross-section of the volcano at `y = 400m` and local minimum heights on the volcano *(red points)*. Write an R expression to detect/find the local minimums in the `y400` vector. How far the local minimums are away from the left (zero)? Discuss your results with the *figure b)* (**Hint:** `?diff, ?sign, ?which`)

Q4:(30pts) Find slopes at the locations `y = 400` and `x = seq(5, 855, 10)` meters. What are the location and value of steepest slope *(negative or positive)*? What would be the best place for a climber needs a rest?

Notes:

- Please, email your answers privately to `sezenismail at gmail com`.
- Answers/codes should be in a text file named `hwX_your_full_name.py`.
- For each answer, write a comment line like `# A1` and write your answer below.
- If you dind't understand a question clearly, email your question to ysb801e@googlegroups.com.